CAUCASIANA
Journal on the biodiversity of the Caucasus and the adjacent regions

Caucasiana 4: 173–194 (2025)
DOI: 10.3897/caucasiana.4.e165329

Research Article

Karyotype characteristics and COI gene data of *Chironomus degelenus* Seisebaev, Bakhtin & Siirin, 2001 (Diptera, Chironomidae) from the Eastern Caucasus (Dagestan) and its relationship with the Nearctic species *Chironomus atrella* (Townes, 1945)

Mukhamed Karmokov[1], Azamat Akkizov[2]

1 *Tembotov Institute of Ecology of Mountain territories KBSC, RAS, Nalchik, Russia*
2 *Institute of Biomedical Problems RAS, Center of medico-ecological researches, Nalchik, Russia*
Corresponding author: Mukhamed Karmokov (lacedemon@rambler.ru)

OPEN ACCESS

Academic editor: Levan Mumladze
Received: 17 July 2025
Accepted: 30 September 2025
Published: 8 December 2025

ZooBank: https://zoobank.org/412626AA-4800-41A1-A159-ED409CBF0506

Citation: Karmokov M, Akkizov A (2025) Karyotype characteristics and COI gene data of *Chironomus degelenus* Seisebaev, Bakhtin & Siirin, 2001 (Diptera, Chironomidae) from the Eastern Caucasus (Dagestan) and its relationship with the Nearctic species *Chironomus atrella* (Townes, 1945). Caucasiana 4: 173–194. https://doi.org/10.3897/caucasiana.4.e165329

Copyright: © Karmokov and Akkizov
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Abstract

This study provides new details on the karyotype characteristics, chromosomal polymorphism, and COI gene data of the Dipteran species *Chironomus degelenus* from the Eastern Caucasus. This is the first record of *Ch. degelenus* in the Eastern Caucasus (and Russia in General). The Caucasian population of this species shows eight banding sequences and low polymorphism. We discovered a new banding sequence, degC2, in *Ch. degelenus*. Only arm C showed an inversion polymorphism. There are striking similarities between the polytene chromosome banding patterns of *Ch. degelenus* and *Ch. atrella*. Therefore, we propose including *Ch. degelenus* in the *Ch. atrella* group of closely related species. Phylogenetic analysis of COI gene sequences reveals two distinct entities referred to as "*Ch. atrella*", with a higher-than-average K2P distance of 6.2%. This pattern is attributed to "dark taxa"; genetic information in databases from unknown and taxonomically challenging species. We also discuss the potential link between the emergence of the species *Ch. degelenus* and *Ch. atrella* and Pliocene climate changes.

Key words: mitochondrial DNA, phylogeny, polytene chromosomes

Introduction

Seisebaev et al. (2001) give the first description of *Chironomus degelenus* from a population living in a waterbody on the Degelen mountain massif in the Abai Region, Eastern Kazakhstan. The territory was once part of the Semipalatinsk Nuclear Test Site. This first description included full data and morphological details of the imago, pupa, and larva of the species, along with a description of its karyotype. Following this initial publication, there have been no further published data on this species, nor further records beyond its original type locality.

It is important to note that the authors of the first description provide the name of the species as *Chironomus degelenus* 1 sp.n. We, due to Kiknadze et

al. (2016), hereinafter in the manuscript specify the name of the species as *Ch. degelenus*.

The biggest DNA databases, GenBank and BOLD, do not contain any DNA data, including sequences of the COI gene, on *Ch. degelenus*. Therefore, this study aims to describe of karyotype characteristics, chromosomal polymorphisms, and COI gene data of *Ch. degelenus* from the Eastern Caucasus. In addition, we compare the chromosomal polymorphism characteristics and DNA data on *Ch. degelenus* from the Caucasus with the results of the earlier study.

Material and methods

Fourth-instar larvae of *Ch. degelenus* were used for both the DNA and karyological studies. We collected these larvae from one site in the Republic of Dagestan (Russian Federation): 28.06.22, 41°16.441'N; 47°48.544'E, Dokuzparinsky district, in southern Dagestan. The collection site was a small lake 1.5 km west of Kurush settlement, at ca. 2620 m a.s.l. The lake has a shape close to that of a circle; the maximum depth is about 3 m, and the water mineralization is about 310 ppm. The collection site location is shown in Fig. 1. The geographic division of the Caucasus follows Gvozdetskii (1963).

The head capsule and body of 10 larvae were slide mounted in a Fora-Berlese solution. These specimens have been deposited at the Tembotov Institute of Ecology of Mountain Territories RAS in Nalchik, Russia. We studied the karyotype and chromosomal polymorphism from a further 24 larvae from the Caucasus region.

The occurrence frequencies of different banding sequences and genotypic combinations were calculated by dividing the number of larvae with a particular banding sequence or genotypic combination by the total number of larvae studied. For example, genotypic combination degC1.2 occurred in 3 larvae, so its frequency of occurrence is 3/24, or 0.125.

We fixed the larvae for karyological study in an ethanol-glacial acetic acid solution (3:1). The preparations of the chromosomes were made using the ethanol-orcein technique (see Dyomin and Ilyinskaya 1988; Dyomin and Shobanov 1990). The banding sequences were designated as per the accepted convention, specifying the abbreviated name of the species, the symbol of the chromosome arm and the sequence number, such as degA1, degB1, etc. (Keyl 1962; Wülker and Klötzli 1973).

We performed the identification of chromosome banding sequences for arms A, E, and F using photomaps by Kiknadze et al. (1991, 2016) following the system of Keyl (1962) and chromosome mapping for arms C and D as per Kiknadze et al. (1991, 2016) following the system of Dévai et al. (1989). The chromosome preparations were studied using a Carl Zeiss Axio Imager A2 microscope.

DNA extraction, amplification and sequencing

We used 12 karyologically studied larvae of *Ch. degelenus* for further DNA extraction. DNA was extracted from the larvae and preserved in 96% ethanol using a Diatom™ DNA Prep 100 kit (Izogen Laboratory Ltd., Moscow, Russia) following the manufacturer's protocol. DNA extraction was performed on vacu-

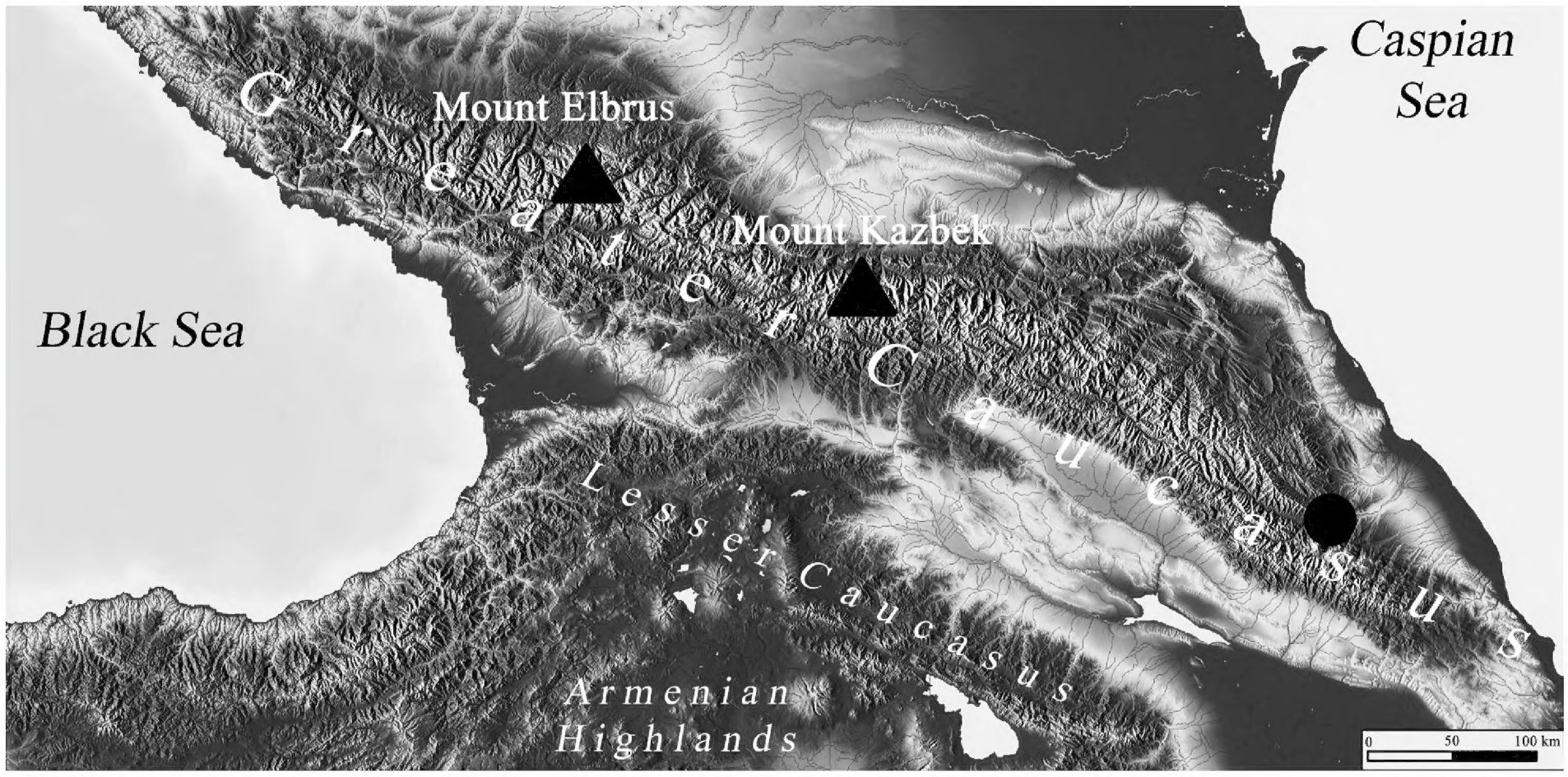

Figure 1. Collection site (dark circle) of *Ch. degelenus* in Eastern Caucasus. The dark triangles marked two highest mountain peaks on the Caucasus ridge. The map was prepared using the web tool available at https://maps-for-free.com/.

um-dried samples without prior homogenization. Samples were incubated in a lysis buffer at a temperature of 55.5 °C for 18 h. After the extraction, the head capsules were retrieved for dry mounting. The resulting DNA solutions were stored at -18 °C. The amplification of the mitochondrial COI gene was conducted using the MasterMix X5 kit (Dialat Ltd., Moscow).

To amplify the mitochondrial COI gene's barcoding region, primers LCO1490 (forward) (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (reverse) (5'- TAAACTTCAGGGTGACCAAAAAATCA-3') (Folmer et al. 1994) were used. PCR was performed in a 25-μL reaction volume. The amplification profile consisted of an initial step of 95 °C for 5 min, followed by 70 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min 50 s, and finally a 10-min extension step at 72 °C, a final elongation at 72 °C (10 min), and final storage at 4 °C. The resulting PCR products were purified by precipitation in a 0.15 M CH3COONa solution in 90% ethanol and then rinsed with 70% ethanol. The results were visualized using 1.5% agarose gel electrophoresis with ethidium bromide.

Purified PCR products were sequenced in both directions. DNA sequencing of the COI gene was performed according to Sanger using the BigDye Terminator v3.1 commercial kit (ThermoFisher) and an ABI 3130xl genetic analyser (ThermoFisher) at Syntol JSC (Moscow, Russia). The GenBank accession numbers of the eight *Ch. degelenus* sequences obtained in this study are OQ873429, OQ873430, OQ873431, OQ873432, OQ873433, OQ873434, OQ873435, and OQ873436.

We used DNA data from both GenBank and BOLD databases (Ratnasingham and Hebert 2007) for the species *Ch. balatonicus* (JN016826.1), *Ch. plumosus* (JN016830.1, AB740262.1, BSCHI661-17, BSCHI063-11, GBDP44180-19), *Ch. agilis* 2 (AF192190.1), *Ch. usenicus* (JN016806.1), *Ch. bonus* (MZ014021.1), *Ch. entis* (GBDPC429-14, KJ085531.1), *Ch. borokensis* (AB74026.1), *Ch. muratensis* (AF192194.1), *Ch. dorsalis* (JN887047.1), *Ch. salinarius* (KC250756.1, KC250745.1), *Ch. annularius* (AF192189.1), *Ch. bernensis* (AF192188.1), *Ch.*

commutatus (AF192187.1), *Ch. atrella* (USDIQ1762-10, USDIQ1587-10, BB-DEE080-10, MHCOL377-07, DRYAS21708-15, SMTPM4256-15), *Ch. athalassicus* (KM995734.1), *Ch. hyperboreus* (HQ941615.1), *Ch. heteropilicornis* (MZ450155.1, MK795770.1, CHMNO266-15), *Ch. pilicornis* (ARCHR026-11, CNQUF171-14), *Ch. staegeri* (KR754417.1), *Ch. frommeri* (KF278235.1), *Ch. acutiventris* (AF192200.1), *Ch. heterodentatus* (AF192199.1), *Ch. melanescens* (AF192204.1), *Ch. sokolovae* (MW471100.1), *Ch. luridus* (AF192203.1), *Ch. pseudothummi* (AF192205.1), *Ch. riparius* (LN894974.1), *Ch. piger* (AF192202.1) and *Pagastiella orophila* (JN265047.1).

The alignment of COI sequences was conducted using MUSCLE with a genetic code for "invertebrate mitochondrial" packaged in MEGA 6 (Tamura et al. 2013). The pairwise sequence distances (Table 2), consisting of the estimated number of base substitutions per site, were calculated using MEGA 6 and the K2P model (Kimura 1980). The analysis involved 51 nucleotide sequences. The codon positions included were the 1st, 2nd, 3rd, and Noncoding. All positions containing gaps and missing data were eliminated. The final data set contained a total of 579 nucleotide positions.

We conducted the estimation of phylogenetic relationships in BEAST V1.10.4 (Suchard et al. 2018) by the Bayesian Markov-chain Monte-Carlo (MCMC) method, using the GTR+G+I substitution model as selected in MEGA 6. The determination of the appropriate model was performed in MEGA 6 (Tamura et al. 2013); the strict clock was used as a clock model and the Yule process as a speciation model. We ran MCMC for 10.000.000 iterations and 1000 iterations of burn-in. Our analysis involved 51 nucleotide sequences, and we eliminated all positions with less than 95% site coverage. There were 579 nucleotide positions in the final data set. We used the COI sequence of *P. orophila* (Genbank accession number JN265047.1) as an outgroup.

We also tried to get average estimates of divergence time between different branches and clusters that appear on the obtained phylogenetic tree (Figs. 3, 4). The age of the most recent common ancestors (TMRCAs) for DNA clades was estimated in BEAST V1.10.4 (Suchard et al. 2018) by the MCMC method, using the GTR+G+I substitution model as selected in MEGA 6. We used a strict clock as a clock model and a constant size as a coalescent model, with the calibration point assumed by Cranston et al. (2012). The time estimate of 36 million years ago (Mya) for the root node of the divergence between *P. orophila* and all *Chironomus* species was used as a first calibration point. Also, we use a time estimate of 5.75–3.43 (standard deviations 1.25 and 0.76, respectively) Mya as a second calibration point for the node of the divergence of the species in the *Ch. plumosus* group of sibling species (Karmokov 2022). We ran MCMC for 10.000.000 iterations and 1000 iterations of burn-in. Tracer v1.7.1 was used to examine the BEAST log file and ESSs for each parameter, which were all > 200.

According to recent studies (Jamnongluk et al. 2003; Stevens et al. 2006), the COI gene sequence diverges at rates ranging from 1.5% to 2.3% per 1 Mya in insects. Under the proposed partitioning scheme and substitution model chosen using Bayes factors, Papadopoulou et al. (2010) found a divergence rate for the COI gene in the study of tenebrionid beetles of 3.54% per 1 Mya (2.69% when paired with the 16S rRNA gene). In this investigation, we used the three frequently accepted mutation rates of 1.5%, 2.3%, and 3.54% to calculate the

divergence time. We determined TMRCAs for the nodes 1 through to 5 of the phylogenetic tree (Fig. 3). The results are presented in Table 3.

Results

We attributed the larvae of *Chironomus* at the studied site to *Ch. degelenus* based on both morphological and chromosomal characteristics. The morphological larval characters of *Ch. degelenus* from the Caucasian site in general are similar to those previously described for this species by Seisebaev et al. (2001). The larva of *Ch. degelenus* is 14–18 mm long, red, and belongs to the plumosus-type. The segment VII bears lateral tubuli, while segment VIII has two pairs of strongly twisted ventral tubuli that extend beyond the posterior parapods. The wrinkled structure on the outer surface of the ventromental plates is absent. The head capsule is yellow, and the occipital sclerite is darkened. The gular region is slightly darkened. The frontoclypeus is pale. The height of the fourth lateral tooth of the mentum is approximately equal to that of the fifth tooth. The third inner tooth of the mandible is yellow and partially to completely fused to the lower margin. The only difference is that in *Ch. degelenus* from the original description, the antennal blade does not reach the fifth segment of the antenna, whereas in the Caucasian population, the antennal blade extends beyond the fifth segment. Morphological characteristics of larvae are presented in Fig. 2A–F.

Karyotype of *Ch. degelenus* from the Eastern Caucasus.

The diploid number of chromosomes in the *Ch. degelenus* karyotype is 2n = 8, and the chromosome arm combinations are AB, CD, EF, and G (the “thummi” cytocomplex) (Fig. 3). Chromosome AB is metacentric, CD and EF are submetacentric, and G is telocentric. The centromeric bands are prominent. There are two nucleoli: one of them is located on the centromeric end of the arm G, and another is in the middle of the arm B. Two Balbiani Rings are developed on the arm G. In arm G, the degree of development of the Balbiani ring and the degree of homologue conjugation in different larvae varied.

Banding sequences and chromosomal polymorphism of *Ch. degelenus* from the Eastern Caucasus.

Previously, Seisebaev et al. (2001) described 7 banding sequences in the *Ch. degelenus* banding sequence pool. In our study, all of those sequences are present, and one new banding sequence, degC2.2, has been found for the first time, providing eight banding sequences in the Caucasian population.

We should note that after we compared the obtained *Ch. degelenus* COI sequences with those in the databases (using the BLAST (The Basic Local Alignment Search Tool) algorithm in the NSBI database), we found that the closest species to *Ch. degelenus* is *Ch. atrella* (Townes, 1945). Looking through the available data in the literature, we see that the karyotypes of these species are surprisingly similar. The patterns of the banding sequences of *Ch. degelenus* have almost identical analogues among the banding sequences of *Ch. atrella*. For our comparison we used the photomaps of *Ch. atrella* from Martin et al.

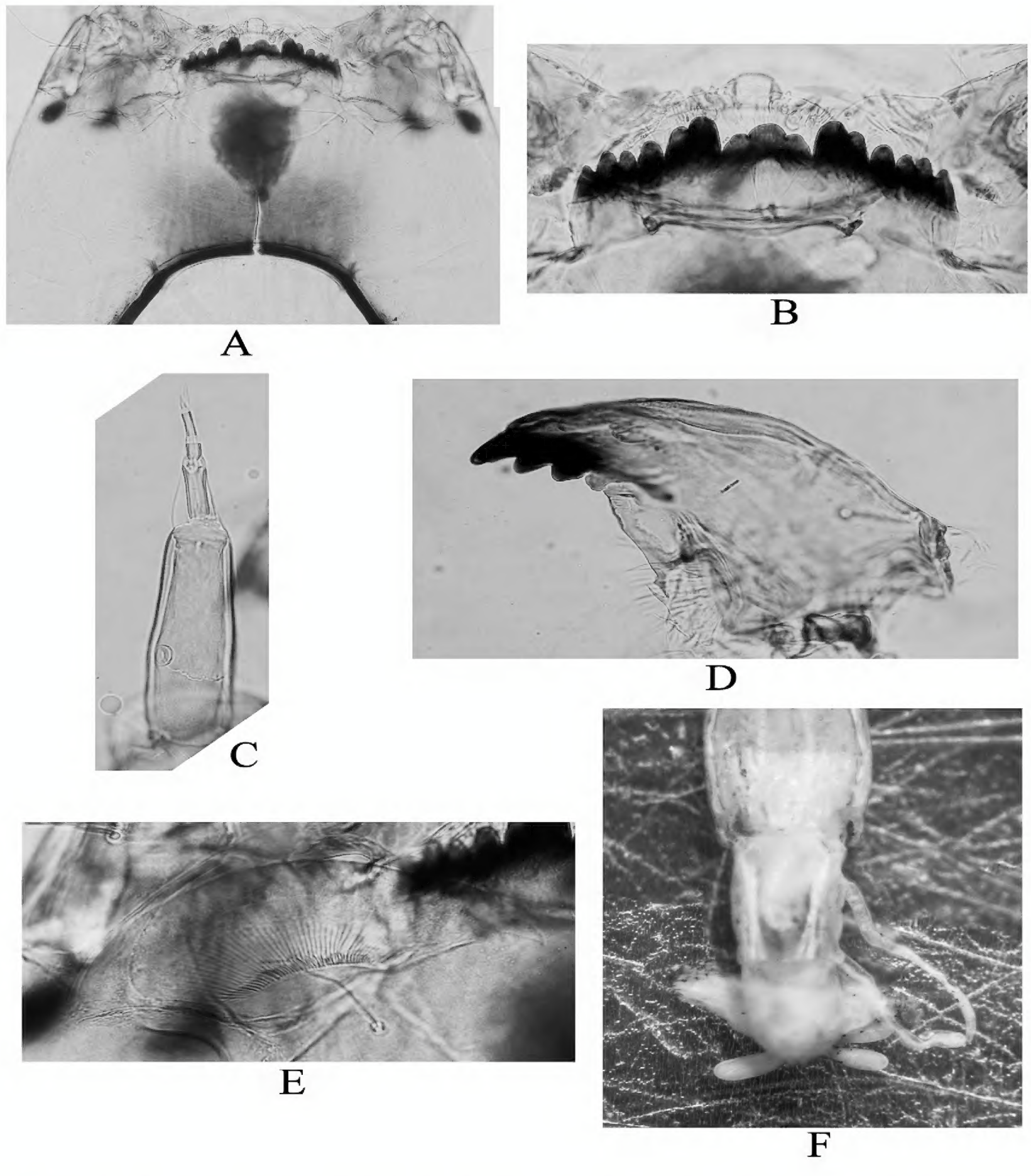

Figure 2. The larva of *Ch. degelenus* from Eastern Caucasus. **A**: gular sclerite; **B**: mentum; **C**: antenna; **D**: mandible; **E**: ventromental plate; **F**: lateral tubuli and ventral tubuli on segment VIII.

(2006). In our opinion, the banding sequences in all chromosome arms, except arm G, are highly similar if not identical. However, the mapping in most parts of the arms is different. To illustrate this, we provide photomaps of *Ch. atrella* with the mapping of Martin et al. (2006) for each arm, along with photomaps of Ch. degelenus with the mapping of Seisebaev et al. (2001) (Fig. 4). The mapping of the original samples by Seisebaev et al. (2001) was followed in the mapping of Caucasian samples.

Arm A has one banding sequence degA1 (Figs 3, 4). The banding patterns of sequence degA1.1 are completely identical to atrA1.1 except for one small feature in the vicinity of the centromere. In other species of *Chironomus*, the centromeric band of this arm is normally represented by a thick band (K), located immediately after the thick-band group 19ef (Keyl, 1962). However, in *Ch. atrel-*

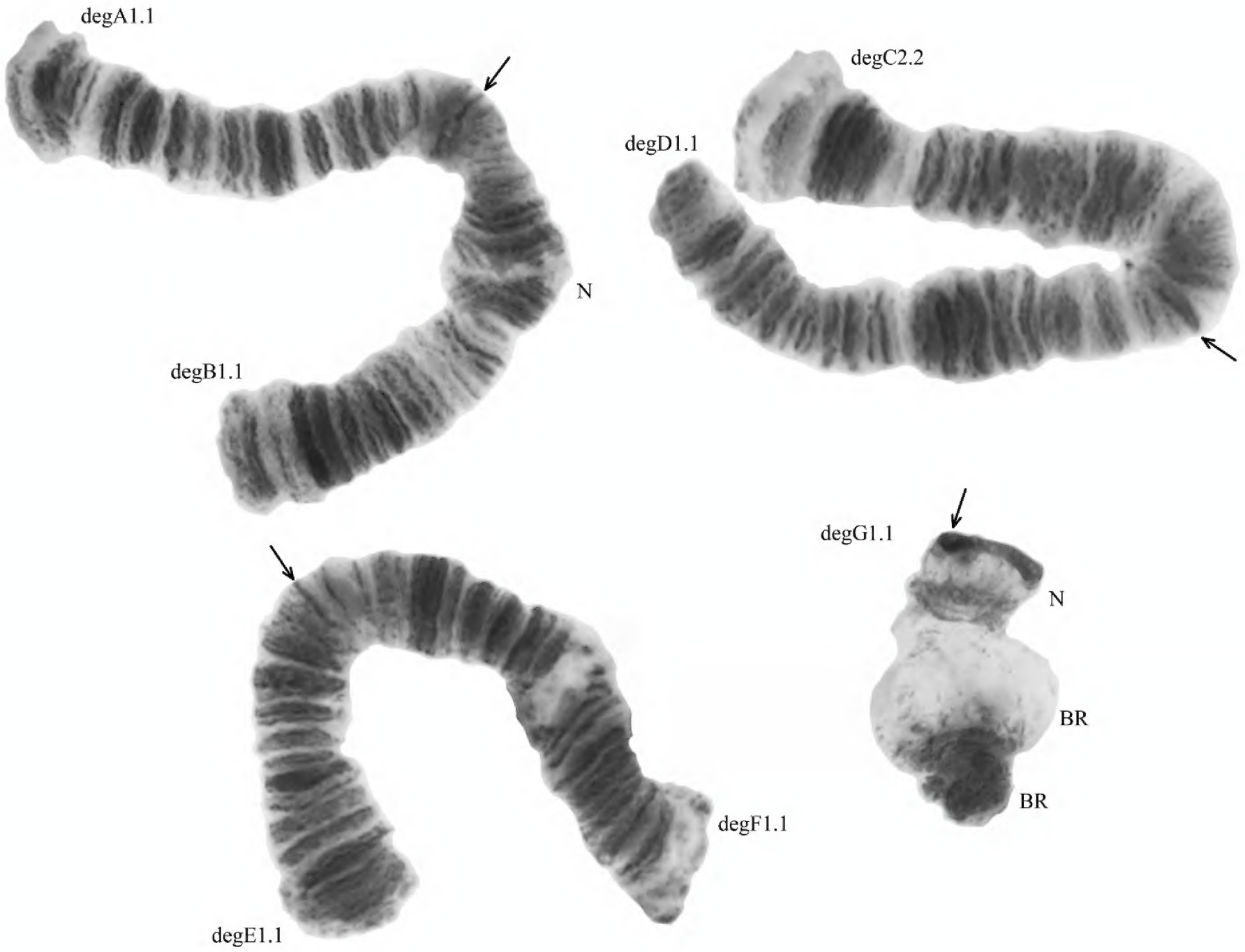

Figure 3. Karyotype of *Ch. degelenus* from the Eastern Caucasus; degA1.1, degB1.1 etc. – genotypic combinations of banding sequences; **BR**: Balbiani rings, **N**: nucleolus. Arrows indicate centromeric bands.

la, there is a significant region containing at least three pale bands between 19ef and K (Martin et al. 2006). The reason for such a structural modification is not clear. Such a region with pale bands is also present in *Ch. degelenus* samples from Kazakhstan but absent in Caucasian ones.

Arm B was monomorphic with banding sequence degB1.1 (Fig. 5). The banding patterns of the sequence is highly similar to banding sequence in arm B that Martin et al. (2001) signed as atrB2+B3.

Arm C has two banding sequences: degC1 and degC2 (Fig. 6). The banding sequence degC2 is new for the species and described for the first time. The banding sequence degC2 (occurrence: 0.708) and genotypic combination degC2.2 (occurrence: 0.542) were predominant in the Eastern Caucasus population o. The sequence degC1 (occurrence: 0.292) and genotypic combination degC1.1 (occurrence: 0.125) were found much less often in the population from the Eastern Caucasus. The heterozygote degC1.2 was also observed with a low frequency (occurrence: 0.333). The sequence degC2 differs from degC1 by simple inversion in the area 8a-11c 4f-a.

The banding patterns of the sequence degC1 (Fig. 6D) is identical to the sequence atrC2 (Fig. 6C) and the sequence degC2 (Fig. 6A) is identical to the sequence atrC1 (Fig 6B).

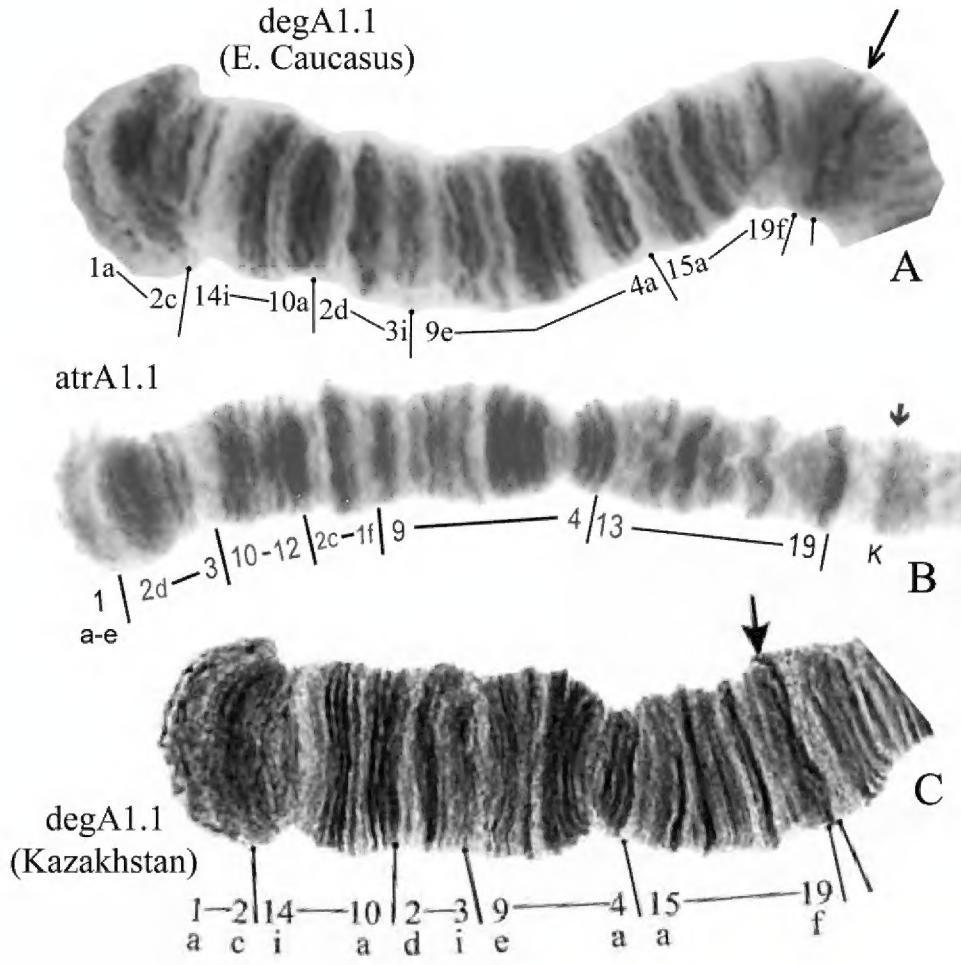

Figure 4. Banding sequences in the arm A of *Ch. degelenus* and *Ch. atrella*. **A**: homozygote degA1.1 from Eastern Caucasus; **B**: homozygote atrA1.1; **C**: homozygote degA1.1 from Kazakhstan from Seisebaev et al., 2001 and Kiknadze et al. (2016). Arrows indicate centromeric bands; **K**: the thick band of the centromeric band. degA1.1 1a-2a-c 14i-10a 2d-3i 9e-4a 15a-19f C.

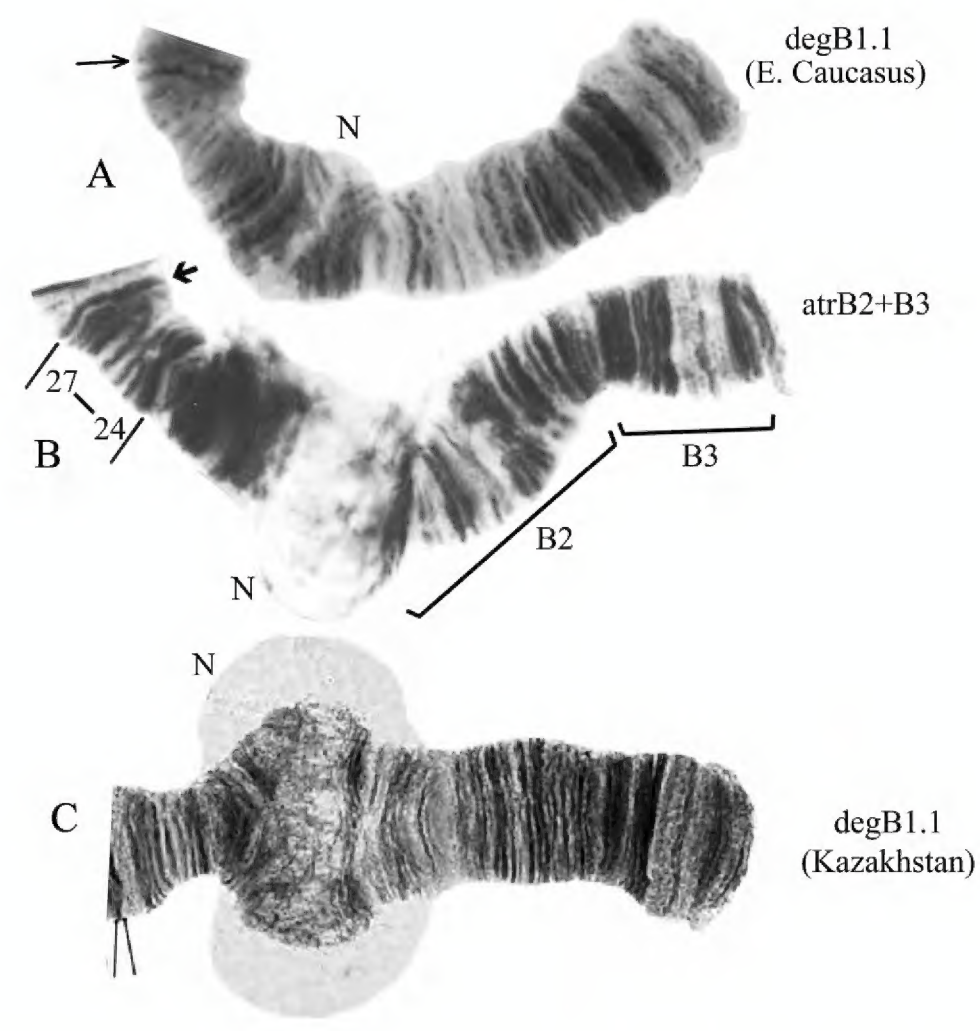

Figure 5. Banding sequences in the arm B of *Ch. degelenus* and *Ch. atrella*. **A**: homozygote degB1.1 from Eastern Caucasus; **B**: homozygote atrB2+B3; **C**: homozygote degB1.1 from Kazakhstan. Designations as in Fig. 3.

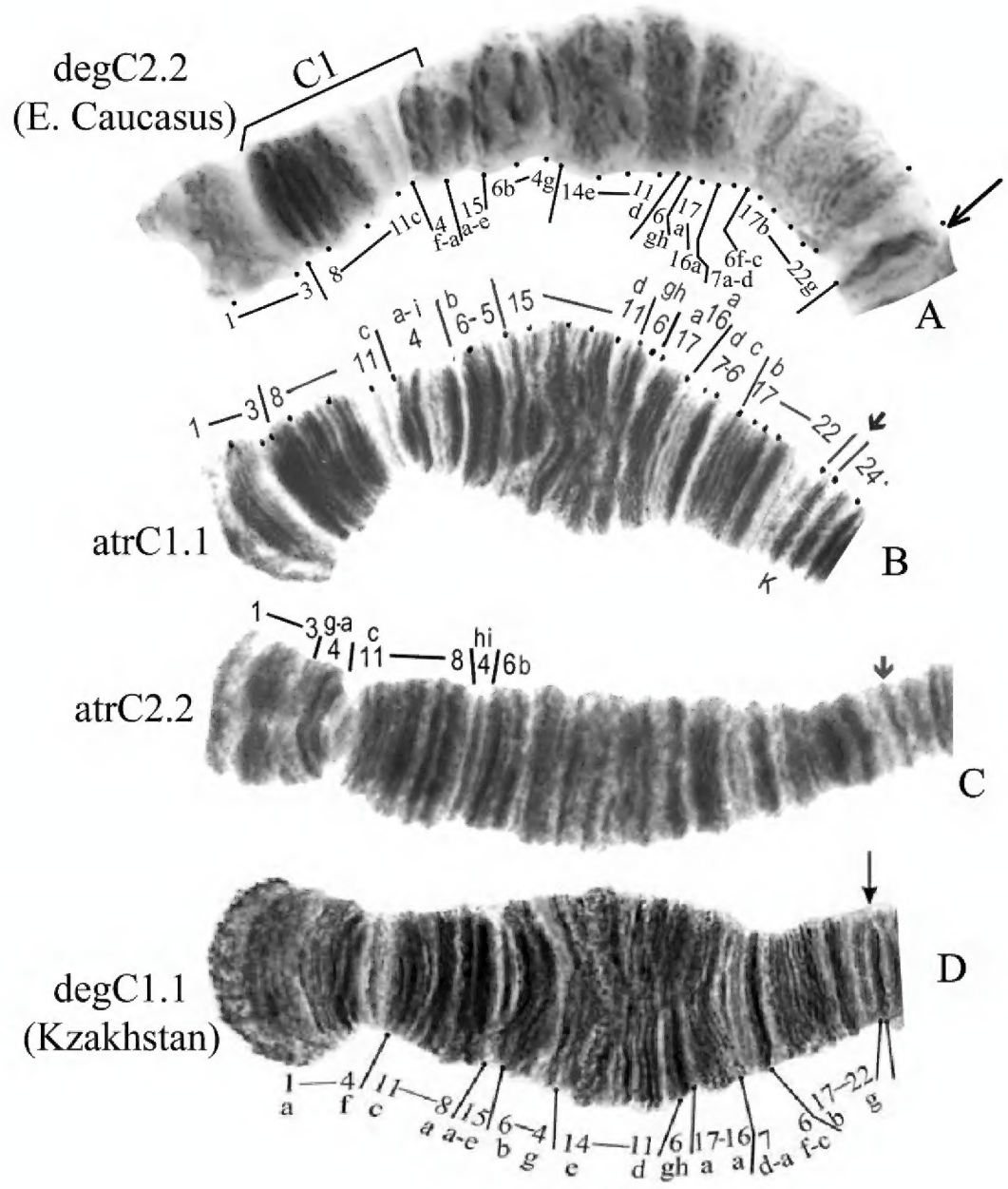

Figure 6. Banding sequences in the arm C of *Ch. degelenus* and *Ch. atrella*. **A**: homozygote degC2.2 from Eastern Caucasus; **B**: homozygote atrC1.1; **C**: homozygote atrC2.2; **D**: homozygote degC1.1 from Kazakhstan. Arrows indicate centromeric bands; **K**: the thick band of the centromeric band. degC1.1 1a-4a-f 11c-8a 15a-e 6b-4g 14e-11h-d 6gh 17a-16h-a 7d-a6f-c 17b-c 18a-22g C; degC2.2 1a-3c 8a-11c 4f-a 15a-e 6b-4g 14e-11h-d 6gh 17a-16h-a 7d-a6f-c 17b-c 18a-22g C.

Arm D also was monomorphic with banding sequence degD1.1 (Fig. 7). The banding patterns of the sequence degD1 (Fig. 7A, C) is identical to the sequence atrD2 (Fig. 7B).

Arm E was monomorphic with banding sequence degE1.1 (Fig. 8). The banding patterns of the sequence degE1 (Fig. 8A, C) is identical to the sequence atrE1 (Fig. 8B). It is the only arm where the chromosome mapping of Martin et al. (2006) for *Ch. atrella* and Seisebaev et al. (2001) for *Ch. degelenus* correspond.

Arm F was monomorphic with a banding sequence degF1.1 (Fig. 9). The banding patterns of the sequence degF1 (Fig. 9A, B, D) are identical to those of the sequence atrF1 (Fig. 9C). Seisebaev et al. (2001) did not note the presence of a nucleolus or Balbiani rings in the arm, but in Caucasian larvae, one can clearly see that in the arm there is some structure resembling a nucleolus. Interestingly, *Ch. atrella* also has a nucleolus in the same region (Martin et al. 2006). The sequence degF1 is still not mapped, while the sequence atrF1 is mapped by Martin et al. (2006), and probably one can use this mapping for sequence degF1.

Arm G was monomorphic with banding sequence degG1.1 (Fig. 3, 10). This is the only arm in the karyotype where chromosome structure is clearly different in both species.

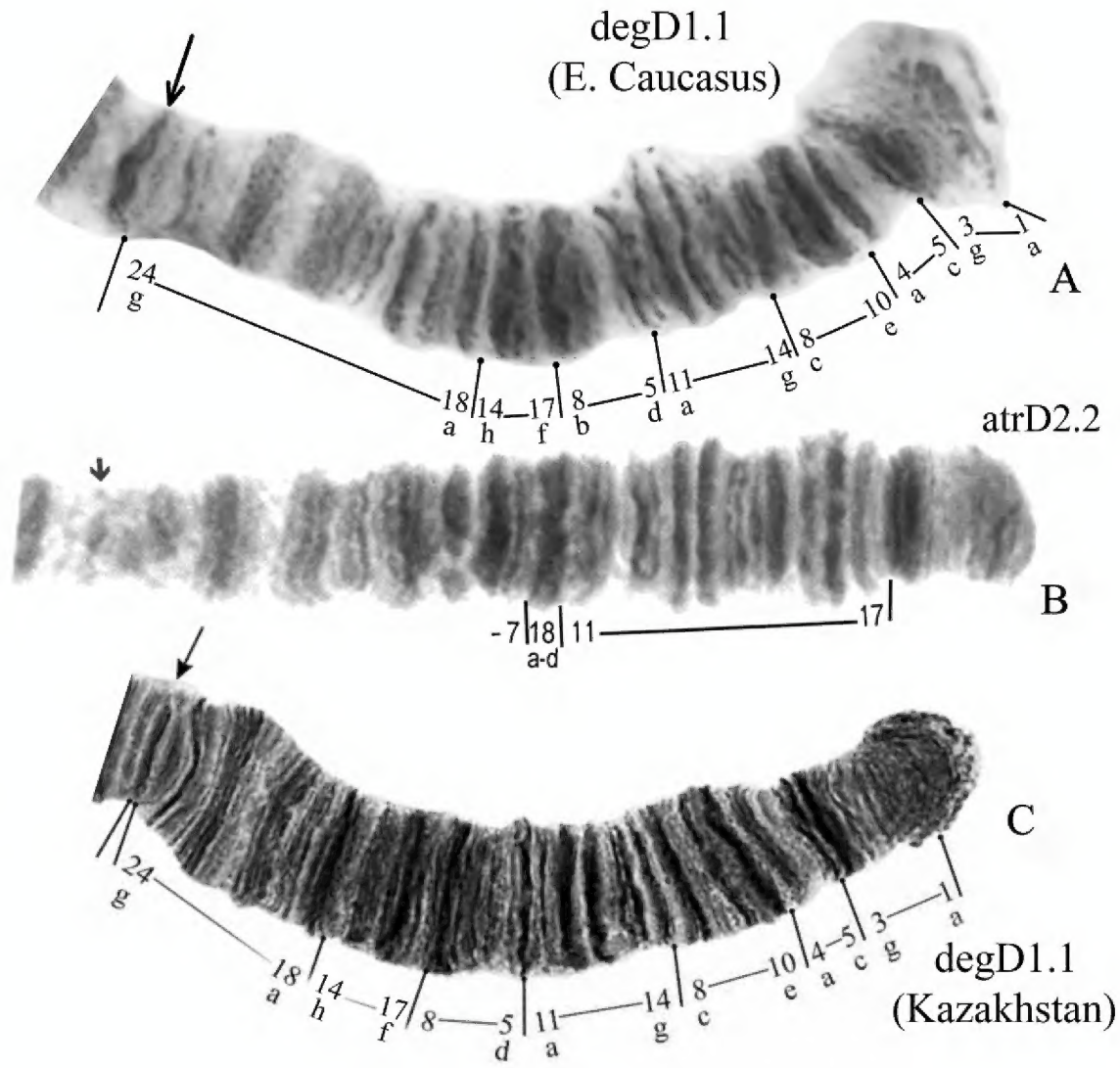

Figure 7. Banding sequences in the arm D of *Ch. degelenus* and *Ch. atrella*. **A**: homozygote degD1.1 from Eastern Caucasus; **B**: homozygote atrD2.2; **C**: homozygote degD1.1 from Kazakhstan. Arrows indicate centromeric bands. deg D1.1 1a-3g 5c-4a 10e-8c 14g-11a 5de-8b 17f-14h 18a-24c C.

The level of inversion polymorphism in the Caucasian *Ch. degelenus* population is quite low. Chromosomal polymorphism was observed only in arm C. The percentage of heterozygous larvae in the population of the Eastern Caucasus is low (33.3%). The average number of heterozygous inversions per larva is also relatively low (0.333).

Results of phylogenetic analysis of COI gene sequences of *Ch. degelenus* and estimated ages of the most recent common ancestors (TMRCAs) for DNA clades.

Overall, we successfully obtained 8 sequences of *Ch. degelenus* from 12 Caucasian larvae. We found six haplotypes among the obtained sequences. Two haplotypes occur twice each among the obtained sequences: the first haplotype: OQ873433, OQ873431, and the second haplotype: OQ873436, OQ873429, respectively. The CG content and base composition of the obtained sequences are provided in Table 1. This is the first DNA data obtained for *Ch. degelenus*.

As we found that the closest related species to *Ch. degelenus* was *Ch. atrella*, we decided to look closer at the *Ch. atrella* DNA data available in the BOLD database. In general, this species is predominantly found in North America. There are a total of 559 COI gene sequences for *Ch. atrella* in the database. They formed two BINs (clusters) with specimens from two countries: Canada (548 sequences) and the USA (11 sequences). The Barcode Index Number System (BINs) in BOLD is an online framework that clusters barcode sequences al-

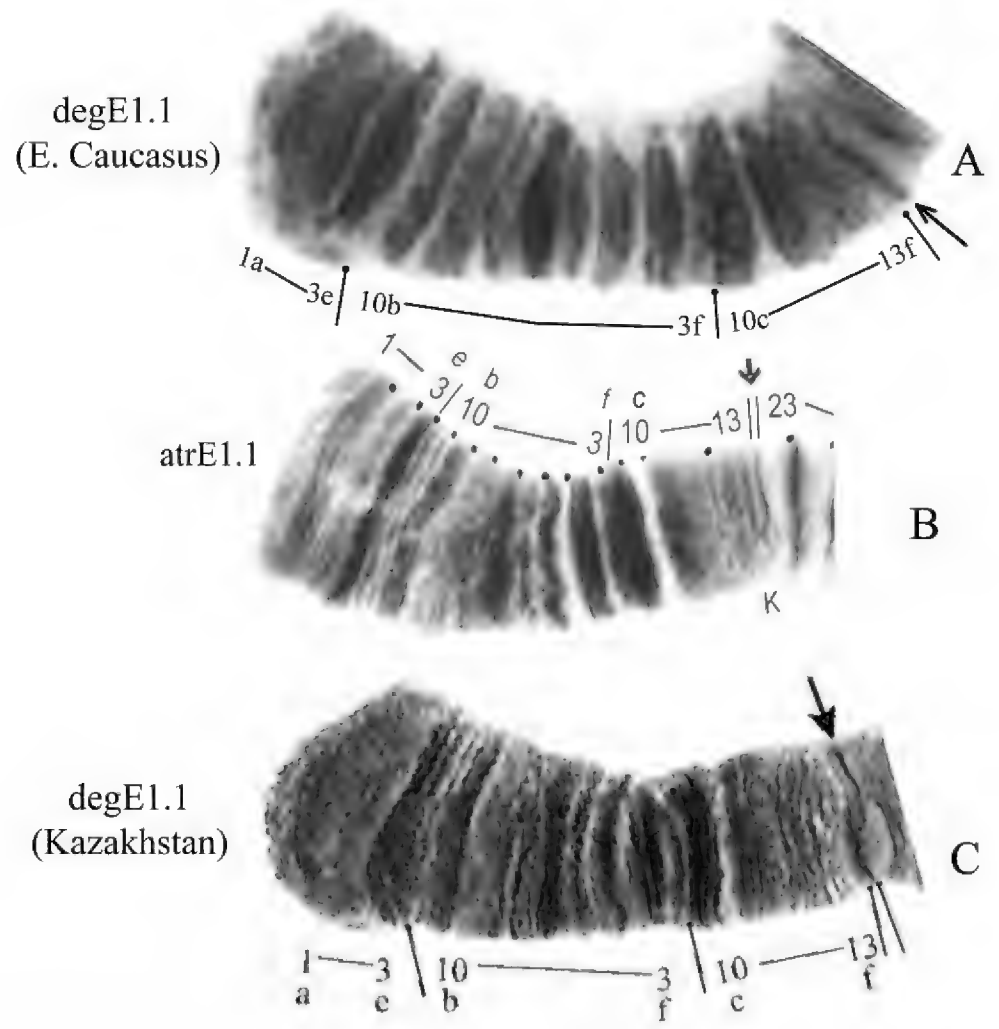

Figure 8. Banding sequences in the arm E of *Ch. degelenus* and *Ch. atrella*. **A**: homozygote degE1.1 from Easetrn Caucasus; **B**: homozygote atrE1.1; **C**: homozygote degE1.1 from Kazakhstan. Arrows indicate centromeric bands; **K**: the thick band of the centromeric band.

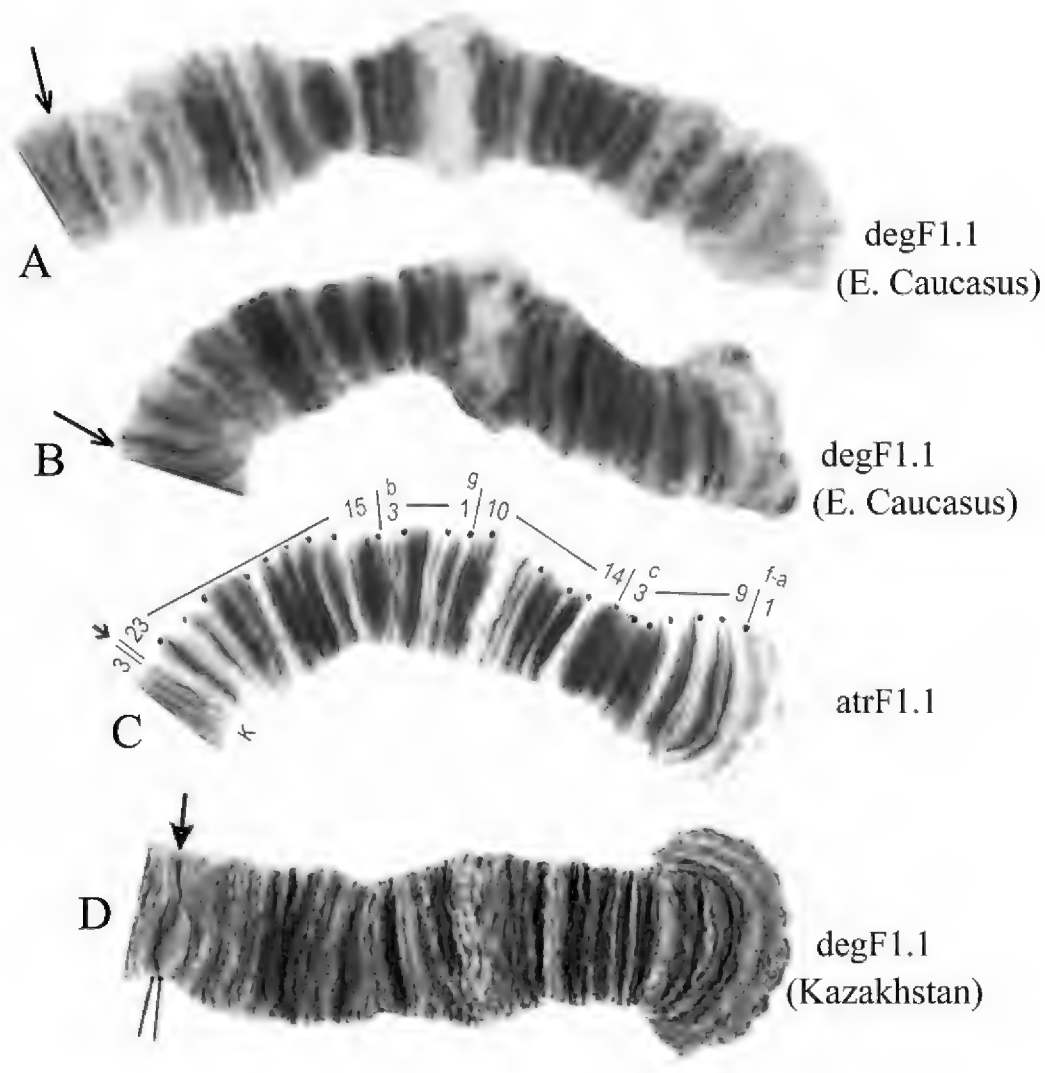

Figure 9. Banding sequences in the arm F of *Ch. degelenus* and *Ch. atrella*. **A, B**: homozygote degF1.1 from Eastern Caucasus; **C**: homozygote atrF1.1; **D**: homozygote degE1.1 from Kazakhstan. Arrows indicate centromeric bands.

Table 1. CG content and base composition in the obtained sequences of *Ch. degelenus*.

Sequences	CG	A	T	G	C
OQ873429 *Ch. degelenus*	37.014%	25.532%	36.018%	17.325%	18.845%
OQ873430 *Ch. degelenus*	36.801%	25.684%	36.170%	17.477%	18.541%
OQ873431 *Ch. degelenus*	36.656%	25.684%	37.082%	17.629%	18.693%
OQ873432 *Ch. degelenus*	36.322%	25.988%	37.690%	17.629%	18.693%
OQ873433 *Ch. degelenus*	36.713%	25.684%	36.930%	17.629%	18.693%
OQ873434 *Ch. degelenus*	36.883%	25.684%	36.474%	17.477%	18.845%
OQ873435 *Ch. degelenus*	36.434%	25.38%	36.930%	17.629%	18.085%
OQ873436 *Ch. degelenus*	36.769%	25.988%	36.474%	17.325%	18.997%

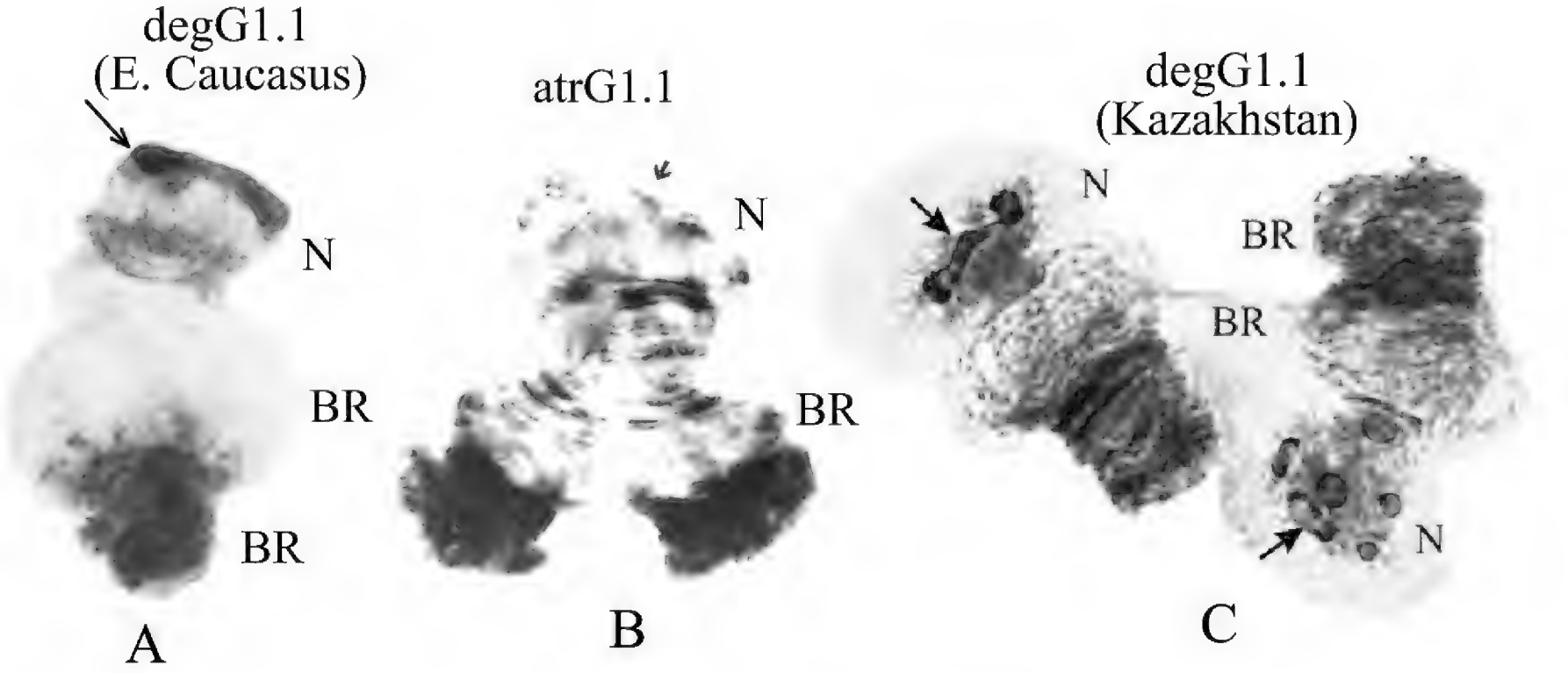

Figure 10. Banding sequences in the arm G of *Ch. degelenus* and *Ch. atrella*. **A**: homozygote degG1.1 from Eastern Caucasus; **B**: homozygote atrG1.1; **C**: homozygote degG1.1 from Kazakhstan. Designations as in Fig. 3.

gorithmically, generating a web page for each cluster. Since clusters show high concordance with species, this system can be used to verify species identifications as well as document diversity when taxonomic information is lacking.

In our analysis, we decided to use the sequences from both BINs and take three haplotypes from each one. The BIN with the largest number of sequences was assigned as the 1st BIN, and the other one as the 2nd BIN. We should note that almost all of the COI sequences available in the BOLD database of *Ch. atrella* were identified by the BOLD Identification System (BOLD-IDS), i.e., through COI gene barcodes without karyological analysis.

The resulting phylogenetic tree (Fig. 11) shows a very complex pattern, each having relatively high probabilities. Some of these clusters correspond well with the groups of closely related species, such as the *Ch. plumosus*, *Ch. pilicornis*, *Ch. piger*, and so on. The most interesting clusters for us are those that contain sequences of *Ch. degelenus* and *Ch. atrella*.

We conditionally named them the *Ch. degelenus* cluster, the *Ch. atrella* 1st BIN cluster, and the *Ch. atrella* 2nd BIN cluster.

The *Ch. degelenus* cluster (Fig. 11) is formed by *Ch. degelenus* sequences from Caucasian specimens. The closest cluster to this is the *Ch. atrella* 1st

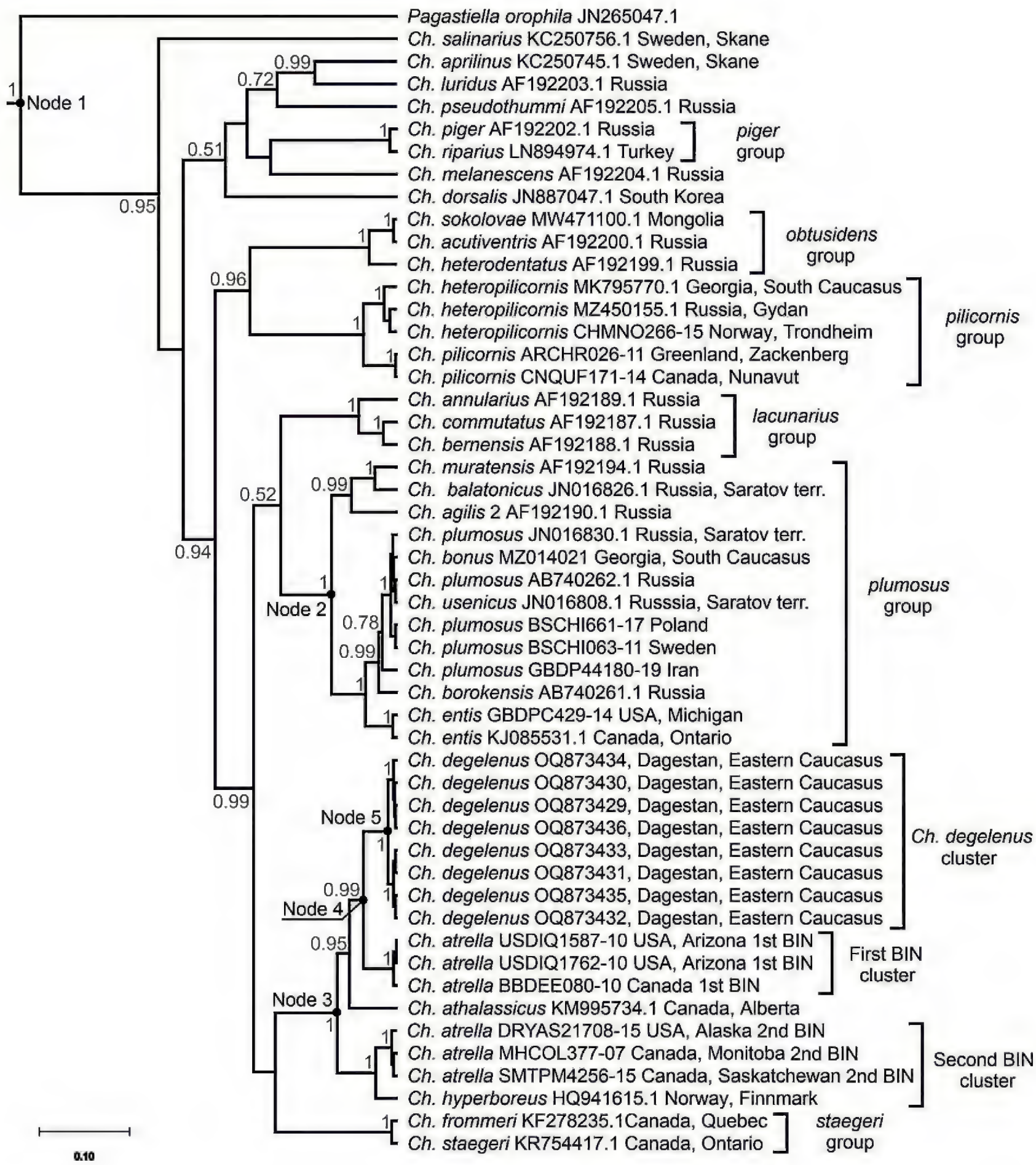

Figure 11. Phylogenetic tree of *Chironomus* species estimated by the Bayesian inference (BI). Support values are given if they exceeded 0.5. The numbers at the nodes indicate posterior probabilities.

BIN cluster, which is formed by sequences of *Ch. atrella* from 1st BIN. Next, the closest to these two clusters is a sequence of *Ch. athalassicus* Cannings, 1975. There are a total of 523 COI gene sequences for *Ch. athalassicus* in the BOLD database. They formed one BIN (cluster) with specimens from two countries: Canada (522 sequences) and the USA (1 sequence).

The sequences of *Ch. atrella* from the 2nd BIN form their own separate cluster, which at some distance is joined by *Ch. hyperboreus* Staeger, 1845. There are a total of 429 COI gene sequences for *Ch. hyperboreus* in the BOLD database. They formed two BINs (clusters) with specimens from three countries: Canada (317 sequences), Greenland (109 sequences), and Norway (3 sequences).

Genetic distances

Calculated pairwise sequence COI gene distances (Fig. 12, Table 2,), consisting of the estimated number of base substitutions per site using the K2P model (Kimura 1980), reveal an interesting pattern. Proulx et al. (2013) demonstrated that intraspecific K2P distances for *Chironomus* species, based on the COI gene ranged from 0% to 3%. They did this by using genetic, morphological, and karyological data to distinguish *Chironomus* species from Canada. These values could potentially be used as a guide in the current study to discriminate between different *Chironomus* species. However, information from the COI gene must be supplemented by other methods.

In the study, the distances between the sequences within the *Ch. degelenus* cluster, the *Ch. atrella* 1st BIN cluster, and the *Ch. atrella* 2nd BIN cluster (Table 2) are below the 3% threshold. The distances within the *Ch. degelenus* cluster are very small, ranging from 0% to 1.65% (average: 0.9%). The distances within the *Ch. atrella* 1st BIN cluster are even smaller, ranging from 0% to 0.36% (average: 0.2%), while those within the *Ch. atrella* 2nd BIN cluster vary from 0% to 0.55% (average: 0.5%).

In contrast, the average distances between the various clusters exceed the 3% threshold. The distance between the *Ch. degelenus* cluster and the *Ch. atrella* 1st BIN cluster is 5.1%, while the distance between the *Ch. degelenus* cluster and the *Ch. atrella* 2nd BIN cluster is 6.6%. Finally, the distance between two *Ch. atrella* clusters is 6.2%.

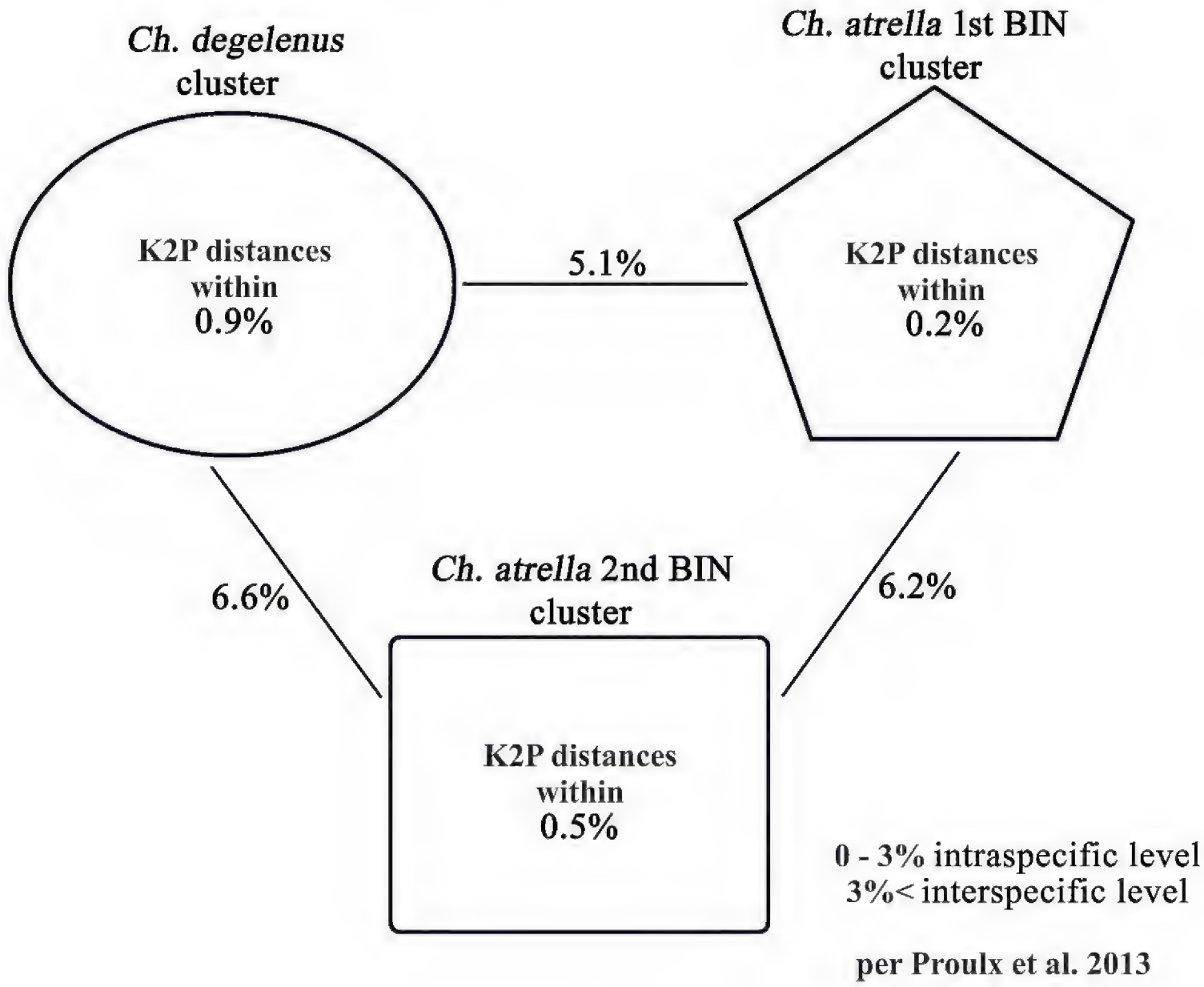

Figure 12. K2P distances within and between *Ch. degelenus* cluster and *Ch. atrella* clusters (1st BIN and 2nd BIN).

Table 2. Estimates of evolutionary divergence between nucleotide sequences of *Ch. degelenus* cluster and *Ch. atrella* clusters (1st and 2nd BINs). The number of base substitutions per site (%) from between nucleotide sequences are shown.

№	Sequences	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	OQ873430 *Ch. degelenus*	0															
2	OQ873432 *Ch. degelenus*	1.65	0														
3	OQ873433 *Ch. degelenus*	1.28	0.36	0													
4	OQ873434 *Ch. degelenus*	0.54	1.65	1.28	0												
5	OQ873435 *Ch. degelenus*	1.47	0.18	0.18	1.46	0											
6	OQ873436 *Ch. degelenus*	0.18	1.47	1.10	0.36	1.28	0										
7	OQ873431 *Ch. degelenus*	1.28	0.36	0.00	1.28	0.18	1.10	0									
8	OQ873429 *Ch. degelenus*	0.18	1.47	1.10	0.36	1.28	0.00	1.10	0								
9	USDIQ1762-10 *Ch. atrella*, USA, Arizona 1st BIN	5.08	5.47	5.08	5.27	5.28	4.88	5.08	4.88	0							
10	USDIQ1587- 10 *Ch. atrella*, USA, Arizona. 1st BIN	5.08	5.47	5.08	5.27	5.28	4.88	5.08	4.88	0.00	0						
11	BBDEE080-10 *Ch. atrella*, Canada, 1st BIN	5.08	5.48	5.09	5.27	5.29	4.89	5.09	4.89	0.36	0.36	0					
12	KM995734.1 *Ch athalassicus*, Canada, Alberta	8.07	7.25	7.26	8.29	7.05	8.08	7.26	8.08	7.28	7.28	7.28	0				
13	MHCOL377-07 *Ch. atrella*, Canada, Monitoba, 2nd BIN	7.06	6.44	6.45	7.45	6.25	7.05	6.45	7.05	6.43	6.43	6.04	7.86	0			
14	DRYAS21708-15 *Ch. atrella*, USA, Alaska, 2nd BIN	6.66	6.04	6.05	7.04	5.85	6.65	6.05	6.65	6.23	6.23	5.84	7.65	0.36	0		
15	SMTPM4256-15 *Ch. atrella*, Canada, Saskatchewa, 2nd BIN	7.06	6.44	6.45	7.45	6.25	7.05	6.45	7.05	6.43	6.43	6.04	7.65	0.55	0.55	0	
16	HQ941615.1 *Ch. hyperboreus*, Norway, Finnmark	7.05	7.04	6.64	7.44	6.85	7.05	6.64	7.05	6.44	6.44	6.04	7.05	2.78	2.40	2.97	0

Table 3. Estimations of the age of the most recent common ancestors (TMR-CAs) for DNA clades.

Number of nodes on the tree	Mean value (Mya)	Stdev.	95% HPD interval	ESS
Divergence rate 1.5%				
node 1	33.031	5.919	22.111. 44.872	5006
node 2	4.160	0.707	2.832. 5.524	7356
node 3	3.721	0.772	2.355. 5.268	5522
node 4	2.045	0.4759	1.192. 3.006	5848
node 5	0.504	0.159	0.216. 0.812	4950
Divergence rate 2.3%				
node 1	29.848	5.366	20.155. 40.785	3863
node 2	3.666	0.616	2.501. 4.879	7587
node 3	3.162	0.646	2.016. 4.466	5259
node 4	1.737	0.403	1.022. 2.535	5246
node 5	0.424	0.134	0.188. 0.685	4461
Divergence rate 3.54%				
node 1	27.233	5.007	17.908. 37.003	4611
node 2	3.225	0.309	2.206. 4.336	7566
node 3	2.667	0.321	1.642. 3.788	5397
node 4	1.465	0.354	0.797. 2.158	5812
node 5	0.356	0.117	0.156. 0.584	4576

Diversification rate

The earliest split between the *Ch. degelenus* cluster and *Ch. atrella* clusters, according to our calculations, occurred between 3.72-2.67 Mya (Fig. 11; Table 3, node 3; substitution rates for all earliest and latest estimates in this chapter are 1.5% and 3.54%, respectively). These estimates fall within the Piacenzian epoch (3.60-2.58 Mya), which is the most recent stage of the Pliocene era (5.33-2.58 Mya). The most recent common ancestor of the *Ch. degelenus* cluster and the *Ch. atrella* 1st BIN cluster lived 2.05 and 1.47 Mya (Fig. 11; Table 3, node 4). This division took place during the Calabrian epoch of the Pleistocene. The most recent common ancestor of *Ch. degelenus* cluster lived between 0.504 and 0.356 Mya (Fig. 11; Table 3, node 5), corresponding to the Chibanian stage of the Pleistocene (0.770-0.126 Mya).

Discussion

This study has discovered the presence of *Ch. degelenus* in the Eastern Caucasus in particular, and Russia as a whole for the first time. The Caucasian population of this species can be characterized as being lowly polymorphic. We found one new banding sequence, degC2, in the banding sequence pool of *Ch. Degelenus*, and observed inversion polymorphism only in arm C.

It appears that the polytene chromosome banding sequences of *Ch. degelenus* and *Ch. atrella* are extremely similar, if not identical. However, based on K2P distances between the clusters of *Ch. degelenus* and *Ch. atrella*, we can still claim that these are still two distinct species, and should be regarded as sibling species. In light of the available data, we suggest including *Ch. degelenus* in the *Ch. atrella* group of related species.

Table 4. Estimations of the age of the most recent common ancestors (TMRCAs) for DNA clades.

Number of nodes on the tree / Divergence rate	1.5%	2.3%	3.54%
node 1	33.031	29.848	27.233
node 2	4.160	3.666	3.225
node 3	3.721	3.162	2.667
node 4	2.045	1.737	1.465
node 5	0.504	0.424	0.356

We can see that there are two distinct entities referred to as "*Ch. atrella*": *Ch. atrella* from the 1st BIN cluster, which has a stronger connection to *Ch. degelenus*, and *Ch. atrella* from the 2nd BIN cluster, which has a stronger connection to *Ch. hyperboreus*. It is unclear which of them is the genuine *Ch. atrella*, although it is probably the one that has a stronger connection to *Ch. degelenus*. The average K2P distance between these two distinct entities is 6.2%, which is significantly higher than the 3% range suggested by Proulx et al. (2013). The most recent common ancestor of these two entities lived a very long time ago, 3.72-2.67 Mya, making it clear that these two distinct entities are in reality separate species. This observation concurs with Martin et al. (2006), who noted that the name *Ch. atrella* has been ascribed to two cytologically distinct species: the one he described in his paper of 2006 and the one reported by Hitchcock and Anderson (1968) from a tidal cove (South Cove) in Old Saybrook, Conn., USA.

Clearly, these observations are one of the numerous manifestations of a phenomenon known as "dark taxa"; the presence in the databases (GenBank and BOLD) of genetic information from unknown and taxonomically challenging species (Page, 2016). Morinière et al. (2019) estimate that of all Chironomidae recorded from Germany, the BOLD database of genetic information contains about 65% of sequences without species-level assignment, or "dark taxa". This phenomenon can greatly complicate the work of various specialists, especially specialists in species of the genus *Chironomus* and closely related genera, where it's very important at the beginning of a piece of work to correctly identify the larval material by karyotype, which will then be used for DNA studies.

We can infer that the most recent common ancestor (Fig. 11, Table 4, node 3) of the *Ch. degelenus* and *Ch. atrella* (both BINs) clusters, as well as the species *Ch. athalassicus* and *Ch. hyperboreus*, lived 3.72-2.67 Mya in the Pliocene epoch. These estimates are relatively close to those we previously found (Karmokov 2022) for the common ancestor of the species in the *Ch. plumosus* group of sibling species (5.75-3.43 Mya). In that paper we explained the species divergence of the group by the dispersal of their common ancestor into newly created habitats that emerged as a result of climate changes in the Pliocene.

Due to the retreat of the forests brought on by the gradual cooling of the climate that started in the earlier Miocene epoch, the Pliocene epoch is distinguished by the appearance of a new type of biome, the first true grasslands. According to Pärtel (2005), true grasslands and Serengeti-like animal grazing communities did not emerge until the Late Miocene in the New World and the Pliocene in the Old World (about 5 Mya).

As a consequence of the heterogeneity of these landscapes, a large number of stagnant water bodies seem to have arisen in various hollows, depressions, and lowlands in the areas cleared of forest. Each of these water bodies could

apparently be characterized by a unique combination of size, shape, depth, temperature profile, mineralization level, etc. Such differences in environmental factors could easily result in different breeding seasons for different populations or individuals, which could eventually result in reproductive isolation and the emergence of new species (Karmokov 2022).

It is also necessary to consider the role of wind in this new type of biome. It is known that chironomids are generally poor fliers that disperse by wind (Armitage et al. 1995) rather than through self-propelled flight within their 'flight boundary layer' (sensu [Taylor 1974]). The active individual flight appears limited to less than 10 km (McLachlan 1983, 1986; McLachlan and Neems, 1996; Krosch et al., 2011) and is restricted by the short duration of the adult life stage, generally just a few days (Oliver 1971; Huryn and Wallace 2000). Service (1997) noted that it seems likely that freshwater mosquitoes breeding in woods and forests are transported less frequently by wind than species breeding in more exposed habitats, such as grasslands, heathlands, open savannas, and deserts. At the same time, Service (1997) provides data on some chironomids (in particular saltwater species) being carried long distances by wind, hundreds of kilometres away from their original habitat.

The free-blowing wind currents in the newly occurring open spaces of the first true grasslands could be an additional positive factor in the process of dispersal and divergence of Chironomids in general and Chironomus species in particular. In the Pliocene epoch this would be the case with *Ch. degelenus*, *Ch. atrella* (both BINs), *Ch. athalassicus*, and *Ch. hyperboreus*.

We propose that divergence in the clusters of *Ch. degelenus* and *Ch. atrella* (both BINs), as well as the species *Ch. athalassicus* and *Ch. hyperboreus*, could have resulted from the dispersal (also with the help of wind) of their common ancestor into these newly formed aquatic habitats of the first true grasslands, just as was probably the case in the common ancestor of the *Ch. plumosus* group of closely related species.

We propose that the common ancestor of *Ch. degelenus*, *Ch. atrella* (both BINs), as well as *Ch. athalassicus* and *Ch. hyperboreus*, originated in the Nearctic realm and was native to this region. Based on the estimated ages of the most recent common ancestors within the clusters of *Ch. degelenus* and *Ch. atrella* (both BINs), as well as *Ch. athalassicus* and *Ch. hyperboreus*, it can be inferred that the distant ancestors of *Ch. degelenus*, which were likely native to the Nearctic, were relatively recent arrivals in the Palaearctic. The ancestral populations of *Ch. degelenus* probably migrated to the Palaearctic through the Beringia land bridge, as the age of the most recent common ancestors in the *Ch. degelenus* cluster is significantly younger compared to those in the *Ch. atrella* (both BIN) clusters, as well as *Ch. athalassicus* and *Ch. hyperboreus*.

Conclusion

In conclusion, the present study has provided valuable insights into the phylogenetic relationships and chromosomal characteristics of *Chironomus* species. The discovery of *Ch. degelenus* in the Eastern Caucasus and Russia for the first time is a notable finding, and the low level of polymorphism observed

in the Caucasian population of the species underscores the need to explore further genetic diversity of the species. Additionally, the identification of a new banding sequence, degC2, in the banding sequence pool of *Ch. degelenus*, adds further evidence of chromosomal diversity in this species.

The observed similarity between the polytene chromosome banding patterns of *Ch. degelenus* and *Ch. atrella* suggests that these two species may be close relatives. However, the analysis of K2P distances between the clusters of *Ch. degelenus* and *Ch. atrella* supports their classification as separate species. The observed picture with the species *Ch. degelenus* and *Ch. atrella* could be seen as one of the numerous manifestations of a phenomenon known as "dark taxa", which can greatly complicate the work of specialists, especially experts in species of the genus *Chironomus* and closely related genera.

We propose that the emergence of grasslands in the Pliocene created new habitats for *Chironomus* species, and this, combined with the role of wind in dispersal, may have been a key factor in the divergence of *Chironomus* species, including *Ch. degelenus*, *Ch. atrella* (both BINs), *Ch. athalassicus*, and *Ch. hyperboreus*. Further research is needed to understand the mechanisms driving speciation in these species and to evaluate the factors responsible for the observed chromosomal divergence.

Overall, this study provides important new insights into the evolutionary history and genetic diversity of *Chironomus* species, shedding light on the complex processes that underlie their divergence and speciation.

Acknowledgements

We would like to thank the anonymous reviewers and editorial board of Caucasiana for their help in improving the manuscript.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

This study was partially supported by a grant № 18-04-00961 from the Russian Foundation for Basic Research (RFBR).

Author contributions

MK: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Visualization; AA: Methodology, Software, Validation, Formal Analysis, Writing – Review & Editing.

Author ORCIDs

Mukhamed Karmokov https://orcid.org/0000-0002-3797-2511
Azamat Akkizov https://orcid.org/0000-0002-8900-7457

Data availability

All of the data that support the findings of this study are available in the main text or Supplementary Information.

References

Armitage PD, Cranston PS, Pinder LCV (1995) The Chironomidae: biology and ecology of non-biting midges. Chapman & Hall, London, 572 pp. https://doi.org/10.1007/978-94-011-0715-0

Cranston PS, Hardy NB, Morse GE (2012) A dated molecular phylogeny for the Chironomidae (Diptera). Systematic Entomology 37(1): 172–188. https://doi.org/10.1111/j.1365-3113.2011.00603.x

Dévai Gy, Miskolczi M, Wülker W (1989) Standardization of chromosome arms B, C and D in *Chironomus* (Diptera, Chironomidae). Acta Biologica Debricina. Supplementum Oecologica Hungarica 2 (1): 79–92.

Dyomin SYu, Iliynskaya NB (1988) Change of the density of polytene chromosomes from different organs of *Chironomus plumosus* larvae. Tsitologiya 30 (4): 407–415. (in Russian)

Dyomin SYu, Shobanov NA (1990) Karyotype of *Chironomus entis* from the *plumosus* group (Diptera, Chironomidae) living in the European part of the Soviet Union. Tsitologiya 32 (10): 1046–1054. (in Russian)

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Gvozdetskii NA (1963) The Caucasus. A Description of Its Nature. State Publishing House of Geographic Literature, Moscow, 261 pp. (in Russian)

Hitchcock SW, Anderson JF (1968) Field-plot tests with insecticides for control of *Chironomus atrella*. Journal of Economic Entomology 61(1): 16–19. https://doi.org/10.1093/jee/61.1.16

Huryn AD, Wallace JB (2000) Life history and production of stream insects. Annual Review of Entomology 45: 83–110.

Jamnongluk W, Baimai V, Kittayapong P (2003) Molecular phylogeny of tephritid fruit flies in the *Bactrocera tau* complex using the mitochondrial COI sequences. Genome 46: 112– 118. https://doi.org/10.1139/g02-113

Karmokov MK (2022) Karyotype characteristics and gene COI sequences of *Chironomus bonus* Shilova et Dzhvarsheishvili, 1974 (Diptera, Chironomidae) from the South Caucasus (Republic of Georgia, Paravani river). Comparative Cytogenetics 16(1): 19–38. https://doi.org/10.3897/compcytogen.v16.i1.79182

Keyl H–G (1962) Chromosomenevolution bei *Chironomus*. II. Chromosomenumbauten und phylogenetische Beziehungen der Arten. Chromosoma 13(4): 464–514. https://doi.org/10.1007/BF00327342

Kiknadze II, Shilova AI, Kerkis IE, Shobanov NA, Zelentsov NI, Grebenyuk LP, Istomina AG, Prasolov VA (1991) Karyoypes and larval morphology in tribe Chironomini. Atlas. Nauka Publish, Novosibirsk, 115 pp. (in Russian)

Kiknadze II, Istomina AV, Golygina VV, Gunderina LI (2016) Karyotypes of Palearctic and Holarctic species of the genus *Chironomus* [Electronic resource]. Russian Academy of Sciences, Siberian Branch, Federal Research Center Institute of Cytology and Genetics. Novosibirsk: Academic Publishing House “GEO”. 489 pp.

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111-120.

Krosch MN, Baker AM, Mather PB, Cranston PS (2011) Spatial population genetic structure reveals strong natal site fidelity in *Echinocladius martini* (Diptera: Chironomidae) in northeast Queensland, Australia. Freshwater Biology. 56(7):1328 – 1341. https://doi.org/10.1111/j.1365-2427.2010.02571.x

Martin J, Andreeva EN, Kiknadze II, Wülker WF (2006) Polytene chromosomes and phylogenetic relationships of *Chironomus atrella* (Diptera: Chironomidae) in North America. Genome. 49(11):1384–92. https://doi.org/10.1139/g06-095

McLachlan AJ (1983) Life history tactics of rain-pool dwellers. Journal of Animal Ecology 52: 545–561.

McLachlan AJ (1986) Sexual dimorphism in midges: strategies for flight in the rainpool dweller *Chironomus imicola* (Diptera: Chironomidae). Journal of Animal Ecology 55: 261–267

McLachlan AJ, Neems RM (1996) Is flight architecture determined by physical constraints or by natural selection? The case of the midge *Chironomus plumosus*. Journal of Zoology 240: 301-308. https://doi.org/10.1111/j.1469-7998.1996.tb05286.x

Morinière J, Balke M, Doczkal D, Geiger MF, Hardulak LA, Haszprunar G, Hausmann A, Hendrich L, Regalado L, Rulik B, Schmidt S, Wägele J-W, Hebert PDN (2019) A DNA barcode library for 5,200 German flies and midges (Insecta: Diptera) and its implications for metabarcoding-based biomonitoring. Molecular Ecology Resources 19(4): 900–928. https://doi.org/10.1111/1755-0998.13022

Oliver DR (1971) Life history of the Chironomidae. Annual Review of Entomology 16: 211–230.

Page RDM (2016) DNA barcoding and taxonomy: Dark taxa and dark texts. Philosophical Transactions of the Royal Society B: Biological Sciences 371: 20150334. http://doi.org/10.1098/rstb.2015.0334

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Molecular Biology and Evolution 27(7): 1659–1672. https://doi.org/10.1093/molbev/msq051

Pärtel M, Bruun HH, Sammul M (2005) Biodiversity in temperate European grasslands: origin and conservation. Grassland Science in Europe 10: 1–14.

Proulx I, Martin J, Carew M, Hare L (2013) Using various lines of evidence to identify *Chironomus* species (Diptera: Chironomidae) in eastern Canadian lakes. Zootaxa 3741: 401–458. https://doi.org/10.11646/zootaxa.3741.4.1

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7: 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Seisebaev AT, Bakhtin MM, Siirin MT (2001) New midge species *Chironomus degelenus* sp. n. (Diptera, Chironomidae). Herald of National Nuclear Centre of Republic Kazakhstan. Radioecology. 3: 123-127. (in Russian)

Service MW (1997) Mosquito (Diptera: Culicidae) dispersal–The long and short of it. Journal of Medical Entomology 34: 579–588. https://doi.org/10.1093/jmedent/34.6.579

Stevens MI, Greenslade P, Hogg ID, Sunnucks P (2006) Southern Hemisphere Springtails: Could Any Have Survived Glaciation of Antarctica? Molecular Biology and Evolution 23(5): 874–882. https://doi.org/10.1093/molbev/msj073

Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution 4(1): vey016. https://doi.org/10.1093/ve/vey016

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729.

Taylor LR (1974). Insect Migration, Flight Periodicity and the Boundary Layer. Journal of Animal Ecology 43(1): 225–238. https://doi.org/10.2307/3169

Wülker W, Klötzli AM (1973) Revision der Gattung *Chironomus* Meig. IV. Arten des lacunarius- (commutatus-) complexes. Archiv für Hydrobiologie 72 (2): 474–489.